

PENGELOMPOKAN SEBARAN TRANSFORMATOR DISTRIBUSI BERDASARKAN KAPASITAS DAYA MENGGUNAKAN METODE NAÏVE BAYES

Studi Kasus: PT. PLN RAYON KOTA SAMARINDA

Rheo Malani¹, Bedi Supratty²

^{1,2} Jurusan Teknologi Informasi, Politeknik Negeri Samarinda, Jl. Cipto Mangunkusumo Kampus Gn Lipan Samarinda – Kalimantan Timur

Email¹: anaogie@gmail.com

Email²: bedirheody@gmail.com

ABSTRAK

Kebutuhan manusia terhadap energi sebagian besar diperoleh dari energi listrik, baik untuk kebutuhan untuk sehari-hari maupun untuk kebutuhan industri. PT. PLN (Persero) salah satu perusahaan listrik negara yang melayani kebutuhan masyarakat akan listrik. Transformator atau lebih dikenal dengan nama “transformator” atau “trafo” sejatinya adalah suatu peralatan listrik yang mengubah daya listrik AC pada satu level tegangan yang satu ke level tegangan berdasarkan prinsip induksi elektromagnetik tanpa merubah frekuensinya. Karena masih kurangnya penyebaran trafo di sekitar wilayah Samarinda dapat berakibat pada layanan kebutuhan listrik kepada masyarakat. Oleh karena itu dibutuhkan suatu metode yang dapat memudahkan penyebaran trafo PT. PLN Rayon Kota Samarinda, salah satu metode tersebut adalah dengan menerapkan Naïve Bayes. Tujuan penelitian ini yaitu memudahkan penyebaran di setiap wilayah dan jenis trafo yang digunakan. Hasil dari perhitungan menggunakan metode Naïve Bayes, diperoleh hasil probabilitas pengelompokan data latih adalah $P(160) = 0.006441224$, $P(100) = 0.016304348$, $P(80) = 0.001610306$, $P(50) = 0.001610306$, $P(40) = 0.000402576$, $P(20) = 0.000679348$. Dari hasil perhitungan, terlihat bahwa nilai probabilitas $P(100)$ lebih dominan, maka 100 direkomendasikan untuk konsumsi real yang digunakan sebagai data latih. Metode Naïve Bayes menghasilkan tingkat akurasi sebesar 92%.

Kata kunci: Transformator, Naïve Bayes, Pengelompokan, Akurasi

ABSTRACT

Human needs for energy are mostly obtained from electrical energy, both for daily needs and for industrial needs. PT. PLN (Persero) is one of the state electricity companies that serves the community's need for electricity. Transformer or better known as "transformer" or "transformer" is actually an electrical device that converts AC power at one voltage level to one voltage level based on the principle of electromagnetic induction without changing its frequency. Because of the lack of distribution of transformers around the Samarinda area, it can result in electricity demand services to the community. Therefore we need a method that can facilitate the distribution of PT. PLN Rayon Kota Samarinda, one of the methods is by applying Naïve Bayes. The purpose of this study is to facilitate the distribution in each region and the type of transformer used. The results of calculations using the Naïve Bayes method, obtained the probability of grouping the training data is $P(160) = 0.006441224$, $P(100) = 0.016304348$, $P(80) = 0.001610306$, $P(50) = 0.001610306$, $P(40) = 0.000402576$, $P(20) = 0.000679348$. From the calculation results, it appears that the probability value $P(100)$ is more dominant, then 100 is recommended for real consumption which is used as training data. The Naïve Bayes method produces an accuracy rate of 92%.

Keywords: Transformer, Naïve Bayes, Grouping, Accuracy

1. PENDAHULUAN

Kebutuhan manusia terhadap energi sebagian besar diperoleh dari energi listrik, baik untuk kebutuhan untuk sehari-hari maupun untuk kebutuhan industri. PT. PLN (Persero) salah satu perusahaan listrik negara yang melayani kebutuhan masyarakat akan listrik. Letak trafo listrik pada PT. PLN pada dasarnya dilakukan dengan cara memenuhi permintaan yang telah ditetapkan oleh kantor pusat untuk dilakukan pemasangan trafo pada setiap wilayah.

Transformator atau lebih dikenal dengan nama “transformer” atau “trafo” sejatinya adalah suatu peralatan listrik yang mengubah daya listrik AC pada satu level tegangan yang satu ke level tegangan berdasarkan prinsip induksi elektromagnetik tanpa merubah frekuensinya. Transformator biasa digunakan untuk mentransformasikan tegangan (menaikkan atau menurunkan tegangan AC). Selain itu, transformator juga dapat digunakan untuk sampling tegangan, sampling arus, dan

juga mentransformasi impedansi. Transformator terdiri dari dua atau lebih kumparan yang membungkus inti besi feromagnetik. Kumparan-kumparan tersebut biasanya satu sama lain tidak dihubungkan secara langsung. Kumparan yang satu dihubungkan dengan sumber listrik AC (kumparan primer) dan kumparan yang lain mensuplai listrik ke beban (kumparan sekunder)[1].

Dalam penyebaran trafo di setiap wilayah harus sesuai kebutuhan wilayah itu sendiri dan dalam pemasangan nya harus menentukan wilayah, berapa trafo, dan jenis trafo yang akan di pasang. Karena masih kurangnya trafo di sekitar wilayah samarinda. Oleh karena itu dibutuhkan suatu metode yang dapat memudahkan penyebaran trafo PT. PLN Rayon Kota Samarinda , salah satu metode tersebut adalah dengan menerapkan algoritma Naïve Bayes.

Peneliti sebelumnya menggunakan algoritma Naïve Bayes dinilai sesuai karena merupakan salah satu algoritma klasifikasi yang sederhana namun memiliki kemampuan dan akurasi tinggi. Aplikasi yang dikembangkan dalam penelitian tersebut yaitu dengan melakukan pengklasifikasian data tweet yang beirisi informasi lalu lintas di kota Bandung. Setelah data terklasifikasikan, kemudian dilakukan visualisasi di peta kota Bandung dengan melalui API Google Map[2].

Metode Naïve Bayes memanfaatkan data training untuk menghasilkan probabilitas setiap kriteria untuk class yang berbeda, sehingga nilai-nilai probabilitas dari kriteria tersebut dapat dioptimalkan untuk memprediksi penggunaan listrik berdasarkan proses klasifikasi yang dilakukan oleh metode Naïve Bayes itu sendiri[3].

Berdasarkan latar belakang, penulis akan meneliti bagaimana pengelompokan penyebaran trafo berdasarkan perhitungan metode naïve bayes sehingga memudahkan penyebaran di setiap wilayah dan jenis trafo yang digunakan.

2. MATERI DAN METODE

A. Clustering

Clustering atau klasterisasi adalah suatu teknik atau metode untuk mengelompokkan data. Menurut Tan, 2006 clustering adalah sebuah proses untuk mengelompokkan data dalam beberapa cluster memiliki tingkat kemiripan yang maksimum dan data antar cluster memiliki kemiripan yang minimum. Clustering merupakan proses partisi satu set objek data ke dalam himpunan bagian yang disebut dengan cluster. Objek yang di dalam cluster memiliki kemiripan karakteristik antar satu sama lainnya dan berbeda dengan cluster yang lain. Partisi tidak dilakukan secara manual melainkan dengan suatu algoritma clustering. Oleh karena itu, clustering sangat berguna dan bisa menemukan group atau kelompok yang tidak dikanal dalam data. Clustering banyak digunakan dalam berbagai aplikasi seperti misalana pada business intelligence, pengenalan pola citra, web search, bidang ilmu biologi, dan untuk keamanan (security). Di dalam business intelligence, clustering bisa mengatur banyak customer ke dalam banyaknya kelompok. Contohnya mengelompokkan customer ke dalam beberapa cluster dengan kesamaan karakteristik yang kuat. Clustering juga dikenal sebagai data segmentasi karena clustering mempartisi banyak data set ke dalam banyak group berdasarkan kesamaannya. Selain itu clustering juga bisa sebagai outlier detection[4][8][9].

B. Metode Naïve Bayes

Naive Bayes merupakan sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan. Algoritma mengunakan teorema Bayes dan mengasumsikan semua atribut independen atau tidak saling ketergantungan yang diberikan oleh nilai pada variabel kelas. Definisi lain mengatakan Naive Bayes merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya[5].

Naive Bayes didasarkan pada asumsi penyederhanaan bahwa nilai atribut secara kondisional saling bebas jika diberikan nilai output. Dengan kata lain, diberikan nilai output, probabilitas mengamati secara bersama adalah produk dari probabilitas individu[6]. Keuntungan penggunaan Naive Bayes adalah bahwa metode ini hanya membutuhkan jumlah data pelatihan yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian.

Persamaan dari teorema Bayes [5] sperti yang ditunjukkan pada persamaan 1.

$$P(H|X) = \frac{P(X|H).P(H)}{P(X)} \quad (1)$$

Di mana :

X : Data dengan class yang belum diketahui

H : Hipotesis data merupakan suatu class spesifik

$P(H|X)$: Probabilitas hipotesis H berdasar kondisi X (posteriori probabilitas)

$P(H)$: Probabilitas hipotesis H (prior probabilitas)

$P(X|H)$: Probabilitas X berdasarkan kondisi pada hipotesis H

$P(X)$: Probabilitas X

Untuk menjelaskan teorema Naive Bayes, perlu diketahui bahwa proses klasifikasi memerlukan sejumlah petunjuk untuk menentukan kelas apa yang cocok bagi sampel yang dianalisis tersebut. Karena itu, teorema bayes di atas disesuaikan dengan persamaan seperti yang ditunjukkan pada persamaan 2.

$$P(C|F_1 \dots F_n) = \frac{P(C).P(F_1 \dots F_n|C)}{P(F_1 \dots F_n)} \quad (2)$$

Dimana Variabel C merepresentasikan kelas, sementara variabel $F_1 \dots F_n$ merepresentasikan karakteristik petunjuk yang dibutuhkan untuk melakukan klasifikasi. Maka rumus tersebut menjelaskan bahwa peluang masuknya sampel karakteristik tertentu dalam kelas C (Posterior) adalah peluang munculnya kelas C (sebelum masuknya sampel tersebut, seringkali disebut prior), dikali dengan peluang kemunculan karakteristik karakteristik sampel pada kelas C (disebut juga likelihood), dibagi dengan peluang kemunculan karakteristik-karakteristik sampel secara global (disebut juga evidence). Karena itu, rumus diatas dapat pula ditulis secara sederhana seperti yang ditunjukkan pada persamaan 3.

$$Posterior = \frac{Prior \times likelihood}{evidence} \quad (3)$$

Dapat dilihat bahwa hasil penjabaran tersebut menyebabkan semakin banyak dan semakin kompleksnya faktor - faktor syarat yang mempengaruhi nilai probabilitas, yang hampir mustahil untuk dianalisa satu persatu. Akibatnya, perhitungan tersebut menjadi sulit untuk dilakukan. Di sinilah digunakan asumsi independensi yang sangat tinggi (naif), bahwa masing-masing petunjuk (F_1, F_2, \dots, F_n) saling bebas (independen) satu sama lain. Dengan asumsi tersebut, maka berlaku suatu kesamaan seperti yang ditunjukkan pada persamaan 4.

$$P(P_i|F_j) = \frac{P(F_i \cap F_j)}{P(F_j)} = \frac{P(F_i).P(F_j)}{P(F_j)} = P(F_i) \quad (4)$$

Untuk $i \neq j$, sehingga

$$P(F_i|C, F_j) = P(F_i|C) \quad (5)$$

Persamaan 5 merupakan model dari teorema Naive Bayes yang selanjutnya akan digunakan dalam proses klasifikasi. Untuk klasifikasi dengan data kontinyu digunakan rumus Densitas Gauss seperti yang ditunjukkan pada persamaan 6.

$$P(X_i = x_i|Y_j = y_j) = \frac{1}{\sqrt{2\pi}\sigma_{ij}} e^{-\frac{(x_i - \mu_{ij})^2}{2\sigma_{ij}^2}} \quad (6)$$

Di mana :

- P : Peluang
- X_i : Atribut ke i
- x_i : Nilai atribut ke i
- Y : Kelas yang dicari
- y_i : Sub kelas Y yang dicari
- μ : mean, menyatakan rata – rata dari seluruh atribut
- σ : Deviasi standar, menyatakan varian dari seluruh atribut.

C. Akurasi

Akurasi merupakan derajat ketepatan antara nilai yang diukur dengan nilai sebenarnya. Nilai replika analisis semakin dekat dengan sampel yang sebenarnya maka semakin akurat metode. Persamaan akurasi dapat dilihat pada persamaan 7.

$$Akurasi = \frac{Jumlah\ data\ benar}{Total\ data\ keseluruhan} \times 100 \quad (7)$$

3. HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Data yang diperlukan dalam penelitian ini diperoleh dari PT. PLN Rayon Kota Samarinda yang berupa data Gardu Rayon Kota Samarinda. Data-data tersebut berisi beberapa kriteria data seperti identitas, alamat, koordinat, pelanggan, konstruksi, kapasitas, konsumsi real, merk, no seri, tahun buat, tap, jurusan, co, la, traves, vooding (mm2), pengamanan utama, NH Out1, NH Out2, NH Out3. Namun dalam penelitian ini hanya beberapa kriteria data saja yang digunakan seperti kapasitas, cadangan, konsumsi real, jumlah konsumsi, dan jumlah konsumen. Dalam penelitian ini penulis menentukan 5 kriteria untuk pengelompokan sebaran trafo. Dari data yang di dapatkan ada 2 kriteria yang sesuai dengan data asli yaitu kapasitas dan konsumsi real. Kemudian penulis memperoleh 3 kriteria yaitu cadangan, jumlah konsumsi, dan jumlah konsumen dari hasil penelitian data-data PT. PLN Rayon Kota Samarinda.

B. Analisa Data

Dari data yang sudah ada kriteria tidak sesuai dengan yang dibutuhkan, maka peneliti menentukan kriteria untuk memenuhi kebutuhan tersebut. Berikut ini hasil kriteria yang di dapatkan.

Tabel 1. Kriteria

Kriteria
Kapasitas
Cadangan
Konsumsi real
Jumlah penggunaan
Penggunaan konsumen

Pada table 1. terdapat 5 kriteria dari sebagian data yang telah disebutkan beberapa data dianalisa oleh peneliti. Analisa data pada penelitian ini berdasarkan data yang dikumpulkan. Data yang diperoleh dari pihak PT PLN yaitu berupa rekap laporan per periode. Pertama yaitu kapasitas didapatkan dari data yang sudah ada. Kedua yaitu cadangan didapatkan informasi dari pihak perusahaan, Ketiga yaitu konsumsi real didapatkan dari data yang sudah ada, Keempat dan Kelima yaitu jumlah penggunaan dan penggunaan konsumsi didapatkan informasi dari pihak perusahaan.

1) Data Latih

Data yang dihasilkan ini berjumlah 23 data. Berikut ini beberapa data yang didapatkan.

Tabel 2. Data Gardu Rayon Kota samarinda

No	Kapasitas (kVA)	Cadangan (kVA)	Konsumsi Real (kVA)	Jumlah Konsumsi (kVA)	Jumlah Konsumen Per Kepala Keluarga
1	200	40	40	80	80
2	200	40	160	125	88
3	100	20	80	62.5	88
4	100	20	40	62.5	32
5	160	50	100	100	88
6	100	20	40	62.5	48
7	200	50	100	125	88
8	200	40	40	125	160
9	25	5	20	15.63	88
10	100	50	20	62.5	52
11	200	40	100	125	88
12	160	50	50	100	96
13	160	32	40	100	88
14	200	40	160	125	120
15	200	50	80	62.5	68
16	160	32	50	125	64
17	160	32	100	100	140
18	200	20	50	62.5	88
19	25	5	20	125	18
20	50	50	40	31.25	18
21	200	20	20	62.5	48
22	100	20	80	125	68
23	250	50	160	156.25	190

2) Data Uji

Adapun data yang akan dikelompokkan berjumlah 12 data. Berikut ini adalah data yang akan di uji.

Tabel 3. Data pengujian

Kapasitas (kVA)	Cadangan (kVA)	Konsumsi Real (kVA)	Jumlah Konsumsi (kVA)	Jumlah Konsumen Per Kepala Keluarga
200	50	100	125	88
100	50	40	100	190
160	32	50	125	64
200	50	100	125	140
160	50	50	62.5	80
100	20	80	125	68
200	40	100	100	88
250	50	160	156.25	88
50	50	40	31.25	18
160	20	0	80	100
100	50	80	62.50	68
200	40	160	125	88

C. Implementasi Metode Naïve Bayes

Dalam bagian ini akan dijelaskan tentang tahapan perhitungan yang dilakukan terhadap data perhitungan pada table II. Berikut ini merupakan langkah-langkah dalam metode Naïve Bayes:

- Menghitung prior probabilities $P(C_i)$ class ke-i / menghitung jumlah class.
- Menghitung posterior probabilities $P(X|C_i)$ setiap atribut dari class ke-i terhadap data X/menghitung jumlah kasus per kelas.
- Menghitung probabilitas hipotesis data/mengalikan semua variable class.
- Menghitung posterior probabilities $P(C_i|X)$ class ke-i terhadap sampel data X/bandingkan hasil perkelas.

Dalam perhitungan ini harus menentukan data latih terlebih dahulu sebelum melakukan perhitungan. Berikut ini data latih yang akan di cari.

Tabel 4. Data Latih

DATA LATIH				
Kapasitas (kVA)	Cadangan	Konsumsi Real	Jumlah Konsumsi	Jumlah Konsumen
200	50	??	125	88

Setelah itu, langkah-langkah metode Naïve bayes yang sudah dipaparkan sebelumnya akan dijabarkan menggunakan data Gardu Rayon Kota Samarinda.

- Pertama, menghitung prior probabilities $P(C_i)$ dari setiap konsumsi real :

$P(\text{Konsumsi Real} = "160") = 3 / 23 = 0.1304$ "jumlah data 160 pada data konsumsi real dibagi dengan jumlah keseluruhan jumlah data.

$P(\text{Konsumsi Real} = "100") = 4 / 23 = 0.1739$ "jumlah data 100 pada data konsumsi real dibagi dengan jumlah keseluruhan jumlah data.

$P(\text{Konsumsi Real} = "80") = 3 / 23 = 0.1304$ "jumlah data 80 pada data konsumsi real dibagi dengan jumlah keseluruhan jumlah data.

$P(\text{Konsumsi Real} = "50") = 3 / 23 = 0.1304$ "jumlah data 50 pada data konsumsi real dibagi dengan jumlah keseluruhan jumlah data.

$P(\text{Konsumsi Real} = "40") = 6 / 23 = 0.2609$ "jumlah data 40 pada data konsumsi real dibagi dengan jumlah keseluruhan jumlah data.

$P(\text{Konsumsi Real} = "20") = 4 / 23 = 0.1739$ "jumlah data 20 pada data konsumsi real dibagi dengan jumlah keseluruhan jumlah data.

Tabel 5. Hasil Perhitungan Jumlah Class $P(C_i)$

$P(C_i)$	Probabilitas
$P(\text{Konsumsi Real} = "160")$	0.1304
$P(\text{Konsumsi Real} = "100")$	0.1739
$P(\text{Konsumsi Real} = "80")$	0.1304
$P(\text{Konsumsi Real} = "50")$	0.1304
$P(\text{Konsumsi Real} = "40")$	0.2609
$P(\text{Konsumsi Real} = "20")$	0.1739

b) Kedua, Menghitung posterior probabilities $P(X|C_i)$ setiap atribut dari class ke-i terhadap data X :

i. *Perhitungan Kapasitas*

$$P(\text{Kapasitas} = "200" | \text{Konsumsi Real} = "160") = 2/3 = 0.6667$$

$$P(\text{Kapasitas} = "200" | \text{Konsumsi Real} = "100") = 2/3 = 0.5000$$

$$P(\text{Kapasitas} = "200" | \text{Konsumsi Real} = "80") = 2/3 = 0.3333$$

$$P(\text{Kapasitas} = "200" | \text{Konsumsi Real} = "50") = 2/3 = 0.3333$$

$$P(\text{Kapasitas} = "200" | \text{Konsumsi Real} = "40") = 2/3 = 2/6 = 0.3333$$

$$P(\text{Kapasitas} = "200" | \text{Konsumsi Real} = "20") = 2/3 = 0.2500$$

Tabel 6. Hasil Perhitungan Kapasitas

$P(X C_i)$	Probabilitas
$P(200 \text{ Konsumsi Real} = "160")$	0.6667
$P(200 \text{ Konsumsi Real} = "100")$	0.5000
$P(200 \text{ Konsumsi Real} = "80")$	0.3333
$P(200 \text{ Konsumsi Real} = "50")$	0.3333
$P(200 \text{ Konsumsi Real} = "40")$	0.3333
$P(200 \text{ Konsumsi Real} = "20")$	0.2500

ii. *Perhitungan Cadangan*

$$P(\text{Cadangan} = "50" | \text{Konsumsi Real} = "160") = 1/3 = 0.3333$$

$$P(\text{Cadangan} = "50" | \text{Konsumsi Real} = "100") = 2/4 = 0.5000$$

$$P(\text{Cadangan} = "50" | \text{Konsumsi Real} = "80") = 1/3 = 0.3333$$

$$P(\text{Cadangan} = "50" | \text{Konsumsi Real} = "50") = 1/3 = 0.3333$$

$$P(\text{Cadangan} = "50" | \text{Konsumsi Real} = "40") = 1/6 = 0.1667$$

$$P(\text{Cadangan} = "50" | \text{Konsumsi Real} = "20") = 1/4 = 0.2500$$

Tabel 7. Hasil Perhitungan Cadangan

$P(X C_i)$	Probabilitas
$P(50 \text{ Konsumsi Real} = "160")$	0.3333
$P(50 \text{ Konsumsi Real} = "100")$	0.5000
$P(50 \text{ Konsumsi Real} = "80")$	0.3333
$P(50 \text{ Konsumsi Real} = "50")$	0.3333
$P(50 \text{ Konsumsi Real} = "40")$	0.1667
$P(50 \text{ Konsumsi Real} = "20")$	0.2500

iii. *Perhitungan Jumlah Konsumsi*

$$P(\text{Jumlah Konsumsi} = "125" | \text{Konsumsi Real} = "160") = 2/3 = 0.6667$$

$$P(\text{Jumlah Konsumsi} = "125" | \text{Konsumsi Real} = "100") = 2/4 = 0.5000$$

$$P(\text{Jumlah Konsumsi} = "125" \mid \text{Konsumsi Real} = "80") = 1/3 = 0.3333$$

$$P(\text{Jumlah Konsumsi} = "125" \mid \text{Konsumsi Real} = "50") = 1/3 = 0.3333$$

$$P(\text{Jumlah Konsumsi} = "125" \mid \text{Konsumsi Real} = "40") = 1/6 = 0.1667$$

$$P(\text{Jumlah Konsumsi} = "125" \mid \text{Konsumsi Real} = "20") = 1/4 = 0.2500$$

Tabel 8. Hasil Perhitungan Jumlah Konsumsi

P(X Ci)	Probabilitas
P(125 Konsumsi Real = "160")	0.6667
P(125 Konsumsi Real = "100")	0.5000
P(125 Konsumsi Real = "80")	0.3333
P(125 Konsumsi Real = "50")	0.3333
P(125 Konsumsi Real = "40")	0.1667
P(125 Konsumsi Real = "20")	0.2500

iv. *Perhitungan Jumlah Konsumen*

$$P(\text{Jumlah Konsumen} = "88" \mid \text{Konsumsi Real} = "160") = 1/4 = 0.3333$$

$$P(\text{Jumlah Konsumen} = "88" \mid \text{Konsumsi Real} = "100") = 1/4 = 0.7500$$

$$P(\text{Jumlah Konsumen} = "88" \mid \text{Konsumsi Real} = "80") = 1/4 = 0.3333$$

$$P(\text{Jumlah Konsumen} = "88" \mid \text{Konsumsi Real} = "50") = 1/4 = 0.3333$$

$$P(\text{Jumlah Konsumen} = "88" \mid \text{Konsumsi Real} = "40") = 1/4 = 0.1667$$

$$P(\text{Jumlah Konsumen} = "88" \mid \text{Konsumsi Real} = "20") = 1/4 = 0.2500$$

Tabel 9. Hasil Perhitungan Jumlah Konsumen

P(X Ci)	Probabilitas
P(88 Konsumsi Real = "160")	0.3333
P(88 Konsumsi Real = "100")	0.7500
P(88 Konsumsi Real = "80")	0.3333
P(88 Konsumsi Real = "50")	0.3333
P(88 Konsumsi Real = "40")	0.1667
P(88 Konsumsi Real = "20")	0.2500

a) Ketiga, Menghitung probabilitas hipotesis data / mengalikan semua variable class :

$$P(X \mid \text{Konsumsi Real} = "160") = P(\text{Kapasitas} = 200, \text{Cadangan} = 50, \text{Jumlah Konsumsi} = 125, \text{Jumlah Konsumen} = 88 \mid \text{Konsumsi Real} = "160") = 2/3 * 1/3 * 2/3 * 1/3 = 0.0494$$

$$P(X \mid \text{Konsumsi Real} = "100") = P(\text{Kapasitas} = 200, \text{Cadangan} = 50, \text{Jumlah Konsumsi} = 125, \text{Jumlah Konsumen} = 88 \mid \text{Konsumsi Real} = "100") = 2/4 * 2/4 * 2/4 * 3/4 = 0.0938$$

$$P(X \mid \text{Konsumsi Real} = "80") = P(\text{Kapasitas} = 200, \text{Cadangan} = 50, \text{Jumlah Konsumsi} = 125, \text{Jumlah Konsumen} = 88 \mid \text{Konsumsi Real} = "80") = 1/3 * 1/3 * 1/3 * 1/3 = 0.0123$$

$$P(X \mid \text{Konsumsi Real} = "50") = P(\text{Kapasitas} = 200, \text{Cadangan} = 50, \text{Jumlah Konsumsi} = 125, \text{Jumlah Konsumen} = 88 \mid \text{Konsumsi Real} = "50") = 1/3 * 1/3 * 1/3 * 1/3 = 0.0123$$

$$P(X \mid \text{Konsumsi Real} = "40") = P(\text{Kapasitas} = 200, \text{Cadangan} = 50, \text{Jumlah Konsumsi} = 125, \text{Jumlah Konsumen} = 88 \mid \text{Konsumsi Real} = "40") = 1/6 * 1/6 * 1/6 * 1/6 = 0.0015$$

$$P(X \mid \text{Konsumsi Real} = "20") = P(\text{Kapasitas} = 200, \text{Cadangan} = 50, \text{Jumlah Konsumsi} = 125, \text{Jumlah Konsumen} = 88 \mid \text{Konsumsi Real} = "20") = 1/4 * 1/4 * 1/4 * 1/4 = 0.0039$$

Tabel 10. Hasil Perhitungan Mengalikan Semua Variabel

$P(X \text{Konsumsi Real}) = 160$	0.0494
$P(X \text{Konsumsi Real}) = 100$	0.0938
$P(X \text{Konsumsi Real}) = 80$	0.0123
$P(X \text{Konsumsi Real}) = 50$	0.0123
$P(X \text{Konsumsi Real}) = 40$	0.0015
$P(X \text{Konsumsi Real}) = 20$	0.0039

b) Keempat, menghitung posterior probabilities $P(X|C_i)$ class ke-i terhadap data konsumsi real :

$$P(X|\text{Konsumsi Real} = "160") * (\text{Konsumsi Real} = "160") = 0.0494 * 3/23 = 0.006441224$$

$$P(X|\text{Konsumsi Real} = "100") * (\text{Konsumsi Real} = "100") = 0.0938 * 4/23 = 0.016304348$$

$$P(X|\text{Konsumsi Real} = "160") * (\text{Konsumsi Real} = "160") = 0.0123 * 3/23 = 0.001610306$$

$$P(X|\text{Konsumsi Real} = "160") * (\text{Konsumsi Real} = "160") = 0.0123 * 3/23 = 0.001610306$$

$$P(X|\text{Konsumsi Real} = "160") * (\text{Konsumsi Real} = "160") = 0.0015 * 6/23 = 0.000402576$$

$$P(X|\text{Konsumsi Real} = "160") * (\text{Konsumsi Real} = "160") = 0.0039 * 2/23 = 0.000679348$$

Tabel 11. Hasil Perbandingan Setiap Class

	0.006441224	160	
	0.016304348	100	
P(X C_i)	0.001610306	80	Hasil
* P(C_i)	0.001610306	50	
	0.000402576	40	
	0.000679348	20	

Dari hasil perhitungan probabilitas masing- masing konsumsi real, ternyata yang paling dominan untuk konsumsi real adalah 100.

D. Pengujian Akurasi

Pengujian ini dilakukan dengan menguji tingkat keakuratan hasil keluaran sistem dengan data dari PT.PLN Rayon Kota Samarinda. Data yang diuji sebanyak 12 data, data Uji yang digunakan terdiri dari 30% data yang sama dengan data latih dan 60% data yang tidak sama dengan data latih. Dari 12 data yang diuji kemudian didapatkan hasil bahwa 1 data memiliki hasil yang tidak sesuai dengan data dari PT.PLN Rayon Kota Samarinda, setelah didapatkan hasil uji coba kemudian dimasukkan ke rumus berikut untuk menghitung tingkat akurasi.

$$Akurasi = \frac{11}{12} \times 100$$

$$Akurasi = 92\%$$

Setelah dihitung dapat disimpulkan bahwa pada penelitian ini metode naïve bayes menghasilkan tingkat akurasi sebesar 92%.

4. KESIMPULAN DAN SARAN

Berdasarkan hasil dari penelitian dengan menggunakan metode Naïve Bayes untuk pengelompokan sebaran trafo, maka di peroleh hasil perhitungan masing-masing dari data konsumsi real, hasil probabilitas pengelompokan data latih adalah $P(160) = 0.006441224$, $P(100) = 0.016304348$, $P(80) = 0.001610306$, $P(50) = 0.001610306$, $P(40) = 0.000402576$, $P(20) = 0.000679348$. Dari hasil perhitungan, terlihat bahwa nilai probabilitas $P(100)$ lebih dominan, maka 100 direkomendasikan untuk konsumsi real yang digunakan sebagai data latih. Metode Naïve Bayes menghasilkan tingkat akurasi sebesar 92%.

Berikut saran untuk penelitian ini terdapat banyak metode untuk pengelompokan, diharapkan untuk para pengembang dapat menggunakan metode tersebut untuk objek yang sama dan membandingkan metode pengelompokan mana yang paling baik dalam kasus pengelompokan tugas akhir untuk pengelompokan sebaran trafo.

DAFTAR PUSTAKA

- [1] A. Gumay, "Transformator," dari <http://eprints.polsri.ac.id>, pp. 1-31, 2015.
- [2] A. Saleh, "Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga," *Citec Journal*, vol. 2, no. 3, pp. 207-217, 2015.
- [3] A. Saleh, "Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga ISSN : 2354-5771," *Citec Journal*, vol. 2, no. 3, pp. 207-217, 2015.
- [4] E. Irwansyah, and M. Faisal, *Advanced Clustering: Teori dan Aplikasi*, p.^pp. 4-5: DeePublish, 2015.
- [5] Bustami, "PENERAPAN ALGORITMA NAIVE BAYES UNTUK MENGLASIFIKASI DATA NASABAH ASURANSI," *Jurnal Informatika*, vol. 8, no. 1, pp. 884-898, 2014.
- [6] M. Ridwan, H. Suyono, and M. Sarosa, "Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier," *JURNAL EECCIS*, vol. 7, no. 1, pp. 59-64, 2013.
- [7] R. G. Rafsanjani, N. Hidayat, and R. K. Dewi, "Diagnosis Penyakit Hati Menggunakan Metode Naive Bayes Dan Certainty Factor," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 2, no. 11, pp. 4478-4482, 2018.
- [8] R. Malani, and B. Suprpty, "Implementasi Sistem Seleksi Kinerja Pegawai Negeri Khususnya Tenaga Pengajar Terbaik Menggunakan Metode K-Means Clustering Pada Politeknik Negeri Samarinda," *Seminar Nasional Teknologi Terapan*, Vol.3, pp.1491, 2016
- [9] P.R. Nastiti, and A.B.W. Putra, "Perbandingan Algoritma K-Means dan Fuzzy C-Means Clustering Untuk Kualifikasi Data Kinerja Dosen di Jurusan Teknologi Informasi POLNES," *Prosiding Seminar SEBATIK*, Vol.1, No.1, pp.71-76, 2017